CLAIMS

What is claimed is:

1. A method of servicing memory requests from at least one memory client, said requests being directed to a memory subsystem having a plurality of independently operable partitions, comprising:

receiving at least one memory request from said at least one client, said memory request including information specifying a location of requested data in the memory subsystem and a data transfer size for said request;

determining one or more partitions needed to service said request based on the specified location and transfer size;

routing said request to each of the one or more servicing partitions; and

servicing the routed request at each servicing partition independently of the other servicing partitions.

2. A method of servicing memory requests from a plurality of memory clients, said requests being directed to a memory subsystem having a plurality of independently operable partitions each with a queue for each of the memory clients, comprising:

receiving a plurality of memory requests from the plurality of clients, each of said memory request including information specifying a location of requested data in the memory subsystem and a data transfer size for said request;

determining for each request, independently of the other requests, one or more partitions needed to service said request based on the specified location and transfer size;

routing to and storing, independently of the other requests, each of said requests in the client queues of the one or more servicing partitions;

selecting, at each of the servicing partitions and independently of other servicing partitions, based on a priority policy, a request from one of the client queues for servicing; and

servicing a request at each servicing partition independently of the other servicing partitions.

3. A method of servicing memory requests from a plurality of memory clients as recited in claim 2, wherein the priority policy is a static policy.

4. A method of servicing memory requests from a plurality of memory clients as recited in claim 3, wherein the static policy is a least recently used policy.

5. A method of servicing memory requests from a plurality of memory clients as recited in claim 3, wherein the static policy is a round-robin policy.

6. A method of servicing memory requests from a plurality of memory clients as recited in claim 3, wherein the static policy is a fixed priority policy.

7. A method of servicing memory requests from a plurality of memory clients as recited in claim 2, wherein the priority policy is a dynamic policy.

8. A method of servicing memory requests from a plurality of memory clients as recited in claim 7, wherein the dynamic policy is one that is selected from the group consisting of a round-robin policy, a fixed priority policy, and a least recently used policy.

9. A method of servicing memory requests from a plurality of memory clients as recited in claim 2, wherein the step of routing to and storing, independently of the other requests, each of said requests in the client queues of the one or more servicing partitions includes performing an address translation of the location specified in said request.

10. A method of servicing memory requests from a plurality of memory clients as recited in claim 2,

further comprising the step of, prior to determining the one or more partitions for each request, selecting one of said requests from a subset of the plurality of clients based on a priority policy; and

wherein the step of determining the one or more partitions for each request includes determining the one or more partitions for the prioritized request.

11. A method of servicing memory requests from a plurality of memory clients as recited in claim 2,

wherein one of the memory clients makes high priority requests; and

further comprising the step of, subsequent to selecting a request from one of the client queues for servicing, selecting one of the selected requests and the high priority requests for servicing by a partition.

12. A method of servicing memory requests from a plurality of memory clients as recited in claim 2,

further comprising the step of, prior to determining the one or more partitions for each request, selecting one of said requests from a subset of the plurality of clients based on a priority policy, wherein each client in the subset has a sub-request ID;

wherein the step of routing to and storing each of the requests in the client queues of the one or more servicing partitions includes storing, in the client queue, the sub-request ID along with the selected request of the subset, the selected request being a read request; and

wherein the step of servicing the read request at each of the servicing partitions includes:

enqueing the sub-request ID together with the read data from the servicing partitions; and

routing the read data to the selected client in the subset based on the enqueued sub-request ID.

13. A method of servicing memory requests from a plurality of memory clients as recited in claim 2,

further comprising the steps of, prior to determining the one or more partitions for each request, selecting one of said requests from a subset of the plurality of clients based on a priority policy, wherein each client in the subset has a sub-request ID;

wherein the step of routing to and storing each of the requests in the client queues of the one or more servicing partitions includes storing, in the client queue, the sub-request ID along with the selected request of the subset, the selected request being a read request; and

wherein the step of servicing the read request at each of the servicing partitions includes:

enqueuing the sub-request ID separately from the read data for the selected client in the subset; and

routing the read data to the selected client in the subset based on the enqueued sub-request ID.

14. A method of servicing memory requests from a plurality of memory clients as recited in claim 2,

wherein at least one of the memory clients is a client that makes requests only to a single partition; and

wherein the step of routing to and storing, independently of the other requests, each of said requests in the client queues of the one or more servicing partitions includes routing to and storing the single-partition requests in the client queue of the single partition.

15. A method of servicing memory requests from a plurality of memory clients as recited in claim 2,

wherein one of the memory requests is a read request; and

wherein the step of servicing a request includes:

determining when a partition servicing the read request has obtained said requested data; and

transferring said requested data from the partition to the memory client.

16. A method of servicing memory requests from a plurality of memory clients as recited in claim 15,

wherein the step of routing to and storing each of the requests in the client queues of the one or more servicing partitions includes appending to the selected read request, sequence count, starting partition, and transfer size information;

wherein the step of servicing a request further includes subsequent to determining when a partition servicing the read request has obtained said requested data, appending to said requested data, sequence count, starting partition, and transfer size information obtained from the selected read request; and

~~wherein the step of transferring said requested data from the partition to the memory client includes accessing the appended sequence count, the starting partition, and the transfer size information to determine when said requested data is available to be sent to the client.~~

~~17. A memory system for servicing memory requests from a plurality of memory clients, comprising:~~

~~a plurality of memory partitions each operable to service a memory request independently of other memory partitions, the memory request including information specifying a location of requested data in the memory subsystem and a data transfer size for said request;~~

~~a plurality of client queues, one for each memory client and each partition;~~

~~a plurality of routing circuits, one routing circuit connected to each memory client and to the client queues, each routing circuit operative to determine for each request, independently of the other requests, one or more partitions needed to service said request based on the specified location and transfer size, each routing circuit operative to route and store, independently of the other requests, each of said requests in the client queues of the one or more servicing partitions; and~~

~~a plurality of arbiter circuits, one arbiter circuit connected to each partition queue and to the client queues for the partition, each arbiter circuit operative to select, independently of other servicing partitions, based on a priority policy, a request for servicing from one of the client queues for the partition, and to transfer each of the selected requests, independently of the other selected requests, to the partition for each of the one or more servicing partitions.~~

~~18. A memory system for servicing memory requests from a plurality of memory clients as recited in claim 17,~~

~~wherein one of said memory requests is a memory read request that requires one or more partitions to provide read data that makes up the read request;~~

~~wherein each partition in the memory subsystem provides a data valid signal that indicates when read data is available from the partition; and~~

~~further comprising:~~

~~a plurality of read queues, one connected to each partition to receive and store said requested read data that is provided by the partition;~~

~~a plurality of queue controllers, each connected to one of the partitions to receive the data valid signal and to one of the read queues to control the read queue, each queue controller producing an indicator of the state of each read queue;~~

~~a side queue that holds data indicating the starting partition and the number of partitions providing read data for the reference;~~

~~multiplexer circuitry having a plurality of inputs, each connected to one of the read queues, the multiplexer circuitry operative to select one of the read queues for output to a particular memory client in response to at least one multiplexer selection control line; and~~

~~a control circuit connected to the side queue to receive data from the side queue, to each of the plurality of queue controllers to receive the status indicators, and to said at least one multiplexer selection control line, the control circuit using said selection control line to select, in response to a status indicator from a queue controller and the data in the side queue, the queue for output to the particular client, starting with the queue for the partition indicated by the side queue.~~

~~19. A memory system for servicing memory requests from a plurality of memory clients as recited in claim 17,~~

~~wherein one of said memory requests is a memory read request that requires one or more partitions to provide read data for said read request;~~

~~further comprising a sequence counter for each memory client, the sequence counter operative to increment a sequence count for each request made by the client to provide a tag for tracking each said read request;~~

~~wherein each partition in the memory subsystem provides a data valid signal that indicates when read data is available from the partition and makes available said sequence count, and data indicating the starting partition and the number of partitions providing read data for the reference with data valid signal and read data from the partition;~~

~~wherein each routing circuit provides a starting partition and burst size information for said read request; and~~

~~further comprising:~~

~~a plurality of read queues, one connected to each partition to receive and store the requested read data that is provided by the partition for said read request and to store with the requested read data the sequence count, the starting partition and burst size for said read request;~~

~~a plurality of queue controllers, each connected to one of the partitions to receive the data valid signal and to one of the read queues to control the read queue, each queue controller producing an indicator of the state of each read queue;~~

~~multiplexer circuitry having a plurality of inputs, each connected to one of the read queues, the multiplexer circuitry operative to select one or more of the read queues for output to a particular memory client in response to at least one selection control line; and~~

~~a control circuit connected to the read queues to receive control data, including said sequence count, data indicating starting partition and number of partitions providing read data for the reference, from the read queues, to each of the plurality of queue controllers to receive the status indicators, and to said at least one selection control line, the control circuit using said selection control line to select, in response to a status indicator from a queue controller and the sequence count, starting partition and burst size in the read queues, the queue for output to the particular client, starting with the queue for the partition indicated by the read queues.~~

~~20. A memory system for servicing memory requests from a plurality of memory clients as recited in claim 18,~~

~~wherein a client accepts data in an order that is non-sequential; and~~

~~wherein the read queue controllers control the read queues to produce data in a non-fifo order to provide read data in the order in which a client accepts the read data.~~

~~21. A memory system for servicing memory requests from a plurality of memory clients as recited in claim 18,~~

~~wherein one of the side queue and read queue contains information indicating that a burst of read data is required of each servicing partition; and~~

~~wherein the multiplexer circuitry includes:~~

~~an output multiplexer for transferring read data to the client; and~~

a plurality of burst registers each connected to one of the read queues and the control circuit, for holding some or all of the burst data in response to the control circuit, wherein the multiplexer circuitry is operative to additionally select one of the burst registers to transfer read data to the output multiplexer.

22. A memory system for servicing memory requests from a plurality of memory clients as recited in claim 18,

wherein there is one partition that services the read request, the servicing partition has a data bus with a particular width for carrying the partition data;

wherein the transfer size specified in the request is smaller that the width of the databus of the servicing partition; and

wherein the multiplexer circuitry includes:

an output multiplexer for transferring read data to the client; and

a plurality of interfacing multiplexers, each connected to one of the read queues and the control circuit, wherein each interfacing multiplexer is operative to sequentially select a first portion and second portion of the databus to send data to the output multiplexer.

23. A memory system for servicing memory requests from a plurality of memory clients as recited in claim 17, wherein the priority policy is a static policy.

24. A memory system for servicing memory requests from a plurality of memory clients as recited in claim 23, wherein the priority policy is a least recently used policy.

25. A memory system for servicing memory requests from a plurality of memory clients as recited in claim 23, wherein the priority policy is a round-robin policy.

26. A memory system for servicing memory requests from a plurality of memory clients as recited in claim 23, wherein the static policy is a fixed priority policy.

27. A memory system for servicing memory requests from a plurality of memory clients as recited in claim 17, wherein the priority policy is a dynamic policy.

~~28. A memory system for servicing memory requests from a plurality of memory clients as recited in claim 17, wherein the routing circuits perform an address translation of the location specified in said request.~~

~~29. A memory system for servicing memory requests from a plurality of memory clients as recited in claim 17, further comprising another arbiter that connects to at least one routing circuit to select one of said requests from a subset of the plurality of clients based on a priority policy and to forward that request to the routing circuit.~~

~~30. A memory system for servicing memory requests from a plurality of memory clients as recited in claim 17,~~

~~wherein one of the memory clients makes high priority requests; and~~

~~further comprising a plurality of secondary arbiters, each connected to the client queue having the high priority request and the arbiter for each of the partitions to select one of the selected requests and the high priority requests for storage in the partition queue of each partition.~~

~~31. A memory system for servicing memory requests from a plurality of memory clients as recited in claim 17,~~

~~wherein at least one of the memory clients is a client that makes requests only to a single partition; and~~

~~wherein the routing circuit routes the single-partition request to the client queue of the single partition.~~

add A